Mechanism of Ubiquitin Activation Revealed by the Structure of a Bacterial MoeB-MoaD Complex

M.W. Lake[1], M.M. Wuebbens[2], K.V. Rajagopalan[2], H. Schindelin[1]

[1]Department of Biochemistry and Center for Structural Biology, State University of New York at Stony Brook

[2]Department of Biochemistry, Duke University Medical Center

The activation of ubiquitin and related protein modifiers[1] is catalyzed by members of the E1 enzyme family, which utilize ATP for the covalent self-attachment of the modifiers to a conserved cysteine. The *Escherichia coli* MoeB and MoaD proteins are involved in molybdenum cofactor (Moco) biosynthesis, an evolutionarily conserved pathway[2]. The MoeB- and E1-catalyzed reactions are mechanistically similar, and despite a lack of sequence similarity, MoaD and ubiquitin display the same fold including a conserved C-terminal Gly-Gly motif[3]. Similar to the E1 enzymes, MoeB activates the C-terminus of MoaD to form an acyl-adenylate. Subsequently, a sulfurtransferase converts the MoaD acyl-adenylate to a thiocarboxylate that acts as the sulfur donor during Moco biosynthesis[4]. These findings suggest that ubiquitin and E1 are derived from two ancestral genes closely related to *moaD* and *moeB*[2]. The crystal structures of the MoeB-MoaD complex in its apo, ATP-bound, and MoaD-adenylate forms presented here highlight the functional similarities between the MoeB- and E1-substrate complexes. These structures provide a molecular framework for understanding the activation of ubiquitin, Rub, SUMO, and the sulfur incorporation step during Moco and thiamine biosynthesis.

The crystal structure of MoeB-MoaD was solved by multiple isomorphous replacement (MIR) using x-ray diffraction data collected at beamline X26C at the National Synchrotron Light Source at Brookhaven National Laboratory. The complex between the *Escherichia coli* MoeB and MoaD proteins reveals a $MoeB_2$-$MoaD_2$ heterotetramer (Fig. 1a) in which the MoeB subunits form a dimer. This dimer interface is primarily hydrophobic and buries a surface area of 5,400 Å^2. To distinguish between the different subunits in the complex, residue numbers are prefixed with either B or D to indicate their location in MoeB or MoaD, respectively.

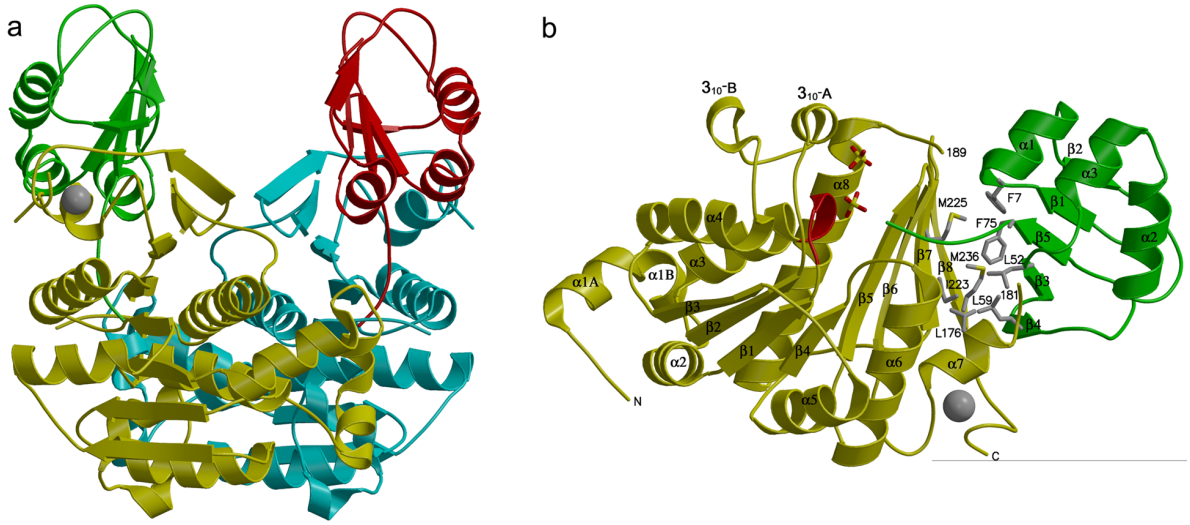

Figure 1. *Structure of the MoeB-MoaD complex.* ***a****) Ribbon diagram of the heterotetramer with MoaD in green and red and MoeB in yellow and cyan. The Zn^{2+}-ions (gray spheres, but one is completely hidden) are coordinated with tetrahedral geometry by four cysteines originating from two Cys-X-X-Cys motifs. All residues (1-81) of MoaD and residues 2-181 and 189-248 of MoeB are observed.* ***b****) Hydrophobic interactions between MoeB (yellow) and MoaD (green). The glycine-rich P-loop motif of MoeB is highlighted in red, and two sulfate molecules bound in the apo-complex are shown in all-bonds representation. One sulfate is ligated by strictly conserved residues in helix 3_{10}-A and the other is in close proximity to the MoaD Gly-Gly motif. Secondary structure elements, terminal residues and those adjacent to the disordered loop are labeled.*

The structure of MoeB consists of eight β-strands that form a continuous sheet surrounded by eight α-helices. In the N-terminal half of the sheet, all β-strands are parallel and reveal a variation of the Rossman fold in which the βαβαβ-topology near the N-terminus is interrupted between the second β-strand and α-helix (β_2 and α_4) by the insertion of two 3_{10} helices. The first 3_{10} helix contains five residues that are strictly conserved in the MoeB/E1 superfamily. A loop between β1 and α3 includes a highly conserved glycine-rich motif (Gly-X-Gly-[Ala/Gly]-[Ile/Leu]-Gly) reminiscent of the P-loop typically found in ATP-hydrolyzing enzymes[5]. The C-terminal half of MoeB contains an antiparallel β-sheet (β5-β8) in a fold distantly related to a family of sugar-binding proteins.

The hydrophobic surface of MoaD involved in MoeB binding is partially conserved in ubiquitin. However, two ubiquitin surface arginines (Arg42 and Arg72) involved in E1 binding[6] are absent from MoaD, indicating that the interactions between ubiquitin and E1 differ to some extent. Perhaps the most striking feature of the MoeB-MoaD interface is the C-terminal extension of residues D76-D81 into a pocket on the MoeB surface. The C-terminus of MoaD extends over β5 of MoeB, which acts as a structural scaffold. Sequence alignments of MoeB and E1 show a preference for small amino acids (Gly, Ala, Ser) at the center of β5, facilitating the insertion of the Gly-Gly motifs of MoaD and ubiquitin into the active sites of MoeB and E1 (Fig. 1b).

In the MoeB-MoaD-ATP ternary complex, ATP is bound in close proximity to the MoaD C-terminus (Fig. 2a) with residues in the P-loop forming the floor of the nucleotide-binding pocket and the adenine ring non-specifically bound in a hydrophobic patch. ATP is anchored at the active site by its triphosphate moiety and ribose hydroxyls. The α-phosphate is buried deeply in the pocket and forms main chain contacts with Gly-B41 of the P-loop. The strictly conserved Arg-B73 contacts one oxygen in each of the α- and β-phosphates. Lys-B86 also interacts with the β-phosphate while Ser-B69 and Asn-B70 anchor the γ-phosphate. The overall shape of the binding pocket distorts the ATP molecule and induces a kink at the α-phosphate. The side chain of Arg-B14' from the second MoeB monomer undergoes a significant conformational change compared to the nucleotide-free structure and is within hydrogen bonding distance of the g-phosphate. Although the MoaD carboxylate and the ATP a-phosphate are in close spatial proximity, nucleophilic attack of the carboxylate appears to be precluded by electrostatic repulsion between the two negatively charged groups.

Soaking experiments with ATP and Mg^{2+} revealed a MoaD acyl-adenylate intermediate consisting of Gly-D81 covalently linked to the α-phosphate through a mixed anhydride (Fig. 2b). Although the pyrophosphate-leaving group is not visible, a bound sulfate molecule from the mother liquor mimics one of the pyrophosphate phosphates. In contrast to glycyl-tRNA synthetase where the metal remains bound to the α-phosphate after formation of the glycyl-adenylate[7], this structure does not contain a bound Mg^{2+}. Comparison of the apo-complex with its ATP-bound and acyl-adenylate forms reveals only subtle conformational changes that are localized to the immediate vicinity of the active site. Other than the conformation of the MoaD C-terminus where Gly-D80 and Gly-D81 adopt clearly different conformations upon acyl-adenylate formation, the active sites of

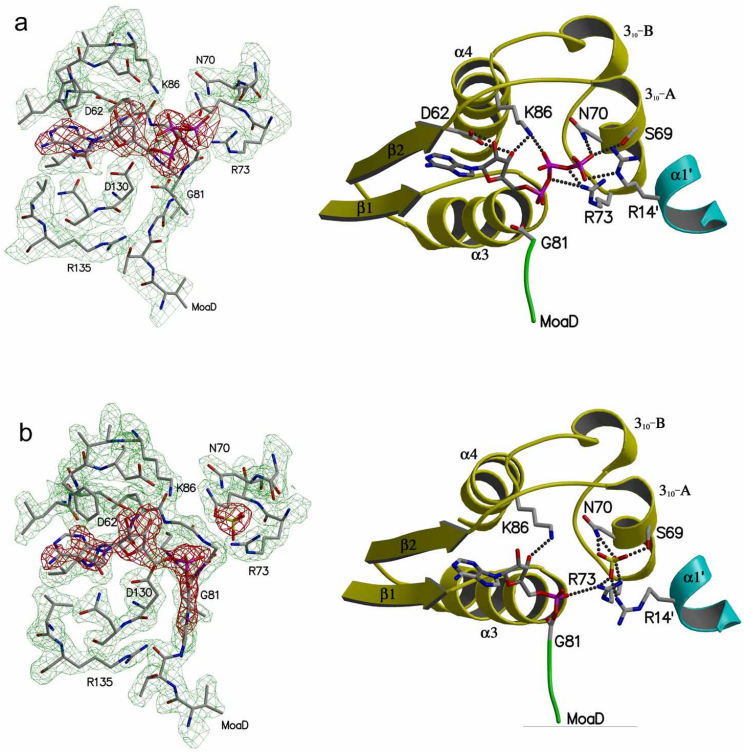

Figure 2. *MoeB catalyzed activation of MoaD.* ***a****) Electron density maps (left) and ribbon diagram (right) of the ATP complex. A $2F_o$-F_c electron density map (green, 1s) encompasses residues from the MoeB active site and MoaD C-terminus. A F_o-F_c electron density map (red, 3σ) shows the ATP. Arg-B14' is shown in cyan in the ribbon diagram.* ***b****) Electron density maps (left), as described in* ***a****, and ribbon diagram (right) of the MoeB-MoaD acyl-adenylate. The difference electron density map covers the covalently bound acyl-adenylate, a sulfate molecule, and Gly-D81.*

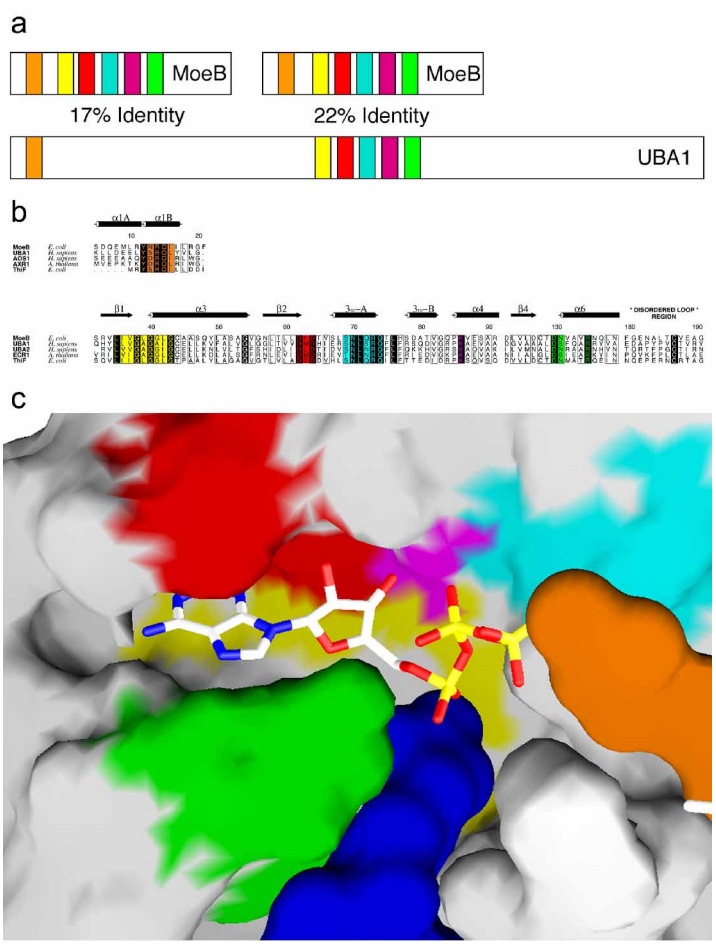

Figure 3. *Active site conservation in the MoeB/E1 enzyme superfamily.* ***a****) Schematic diagram of sequence relationships between MoeB and UBA1, with conserved sequence motifs indicated by colored blocks.* E. coli *MoeB and residues 1 to 250 and 400 to 660 of human UBA1 are 17% and 22% identical, respectively.* ***b****) Excerpts from a multiple sequence alignment of MoeB/E1 superfamily enzymes. The alignment is based on several representatives of each group, but only one family member is displayed: MoeB (*E. coli*), UBA1 (human ubiquitin-activating enzyme), AOS1/UBA2 (heterodimeric human SUMO-activating enzyme), AXR1/ECR1 (heterodimeric* Arabidopsis thaliana *Rub-activating enzyme) and ThiF (*E. coli*). The strictly conserved cysteine corresponding to MoeB Cys187 is also included.* ***c****) Surface representation of the MoeB-MoaD complex around the active site. The surface has been color-coded to correspond to the conserved regions shown in* ***a*** *and* ***b****, with the MoaD C-terminus in blue. The bound ATP molecule is shown in all-bonds representation.*

the apo and acyl-adenylate models are remarkably similar. In contrast, the ATP-bound model shows the most pronounced structural changes, particularly in the side chain of Arg-B14' (Fig. 2a, b).

Multiple sequence alignments of MoeB and different members of the E1 family including the ubiquitin-, RUB- and SUMO-activating enzymes (Fig. 3a, b) reveal a remarkable degree of conservation for the residues surrounding the active site (Fig. 3c). Based on the structural evidence presented here, it is possible to assign functional roles to most of these residues. The loop region between β_1 and α_3 (yellow) consists of a glycine-rich nucleotide-binding motif that facilitates ATP entry into the active site. The loop region between β2 and helix 3_{10}-A (red) is critical for binding the ribose. The highly conserved residues forming helix 3_{10}-A (cyan) are essential for binding the β- and γ-phosphates of ATP and, more importantly, stabilizing the pyrophosphate leaving group upon attack by the MoaD- or ubiquitin-carboxylates. Residues in the loop between β4 and α6 (green) are responsible for proper positioning of Asp-B130 (predicted to be involved in Mg^{2+}-ligation) adjacent to the a-phosphate. In this same region, Arg-B135 inside helix α6 properly orients both the incoming C-terminal extension of MoaD and strand β5 of MoeB, which serves to support the C-terminal MoaD Gly-Gly dipeptide.

In light of the sequence homologies, our studies suggest that enzymes involved in the activation of ubiquitin, Rub, SUMO and ThiS all contain a structurally similar, MoeB-like domain. Careful sequence analysis reveals that the ubiquitin activating enzymes contain an additional MoeB-like domain near their N-terminus (Fig. 3a). While most of the residues in the signature sequence motifs required for ATP-binding and hydrolysis (Asn70, Arg73, Lys86, Asp130 and Arg135) are missing from the first MoeB-like domain, the residue corresponding to Arg14 is strictly conserved. As described above, Arg14 is inserted into the active site across the dimer interface and plays a critical role during ATP hydrolysis. Although the enzymes involved in the activation of SUMO and Rub only contain a single MoeB-like domain, there are additional enzymes in each of these pathways (AOS1 and AXR1) that are essential for catalysis and also contain a MoeB-like domain with an Arg residue corresponding to Arg14 (Fig. 3b). These results strongly suggest that this group of proteins mimics the dimeric structure of MoeB by arranging two MoeB-like domains, present on the same or two different polypeptide chains, in a manner similar to that observed here in the MoeB dimer. In contrast to the two active sites of the MoeB dimer, the monomeric forms of these enzymes are predicted to contain one active site created by residues originating from both MoeB-like domains. The results presented here reveal that while members of the MoeB/E1 enzyme super-

family have diversified to be utilized in seemingly unrelated pathways, their mechanism for acyl-adenylate formation has been evolutionarily conserved.

Acknowledgements

We thank M. J. Rudolph for initial help with crystallization and data collection; J. Daniels for technical assistance; and D. Schneider for support at beamline X26C. This work was supported by NIH Grant DK54835. Research carried out at the National Synchrotron Light Source, Brookhaven National Laboratory, which is supported by the U.S. Department of Energy, Division of Materials Sciences and Division of Chemical Sciences, under Contract No. DE-AC02-98CH10886.

References

1. Hochstrasser, M. Evolution and function of ubiquitin-like protein conjugation systems. *Nature Cell. Biol.* **2**, E153-E157 (2000).
2. Rajagopalan, K. V. Biosynthesis and processing of the molybdenum cofactors. *Biochem. Soc. Trans.* **25**, 757-761 (1997).
3. Rudolph, M. J., Wuebbens, M. M., Rajagopalan, K. V. & Schindelin, H. Crystal structure of molybdopterin synthase and its evolutionary relationship to ubiquitin activation. *Nature Struct. Biol.* **8**, 42-46 (2001).
4. Leimkühler, S., Wuebbens, M. M. & Rajagopalan, K. V. Characterization of *Escherichia coli* MoeB and its involvement in the activation of MPT synthase for the biosynthesis of the molybdenum cofactor. *J. Biol. Chem.* **276,** 34695-34701 (2001).
5. Walker, J. E., Saraste, M., Runswick, M. J. & Gay, N. J. Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J.* **1**, 945-951 (1982).
6. Burch, T. J. & Haas, A. L. Site-directed mutagenesis of ubiquitin. Differential roles for arginine in the interaction with ubiquitin-activating enzyme. *Biochemistry* **33**, 7300-7308 (1994).
7. Arnez, J. G., Dock-Bregeon, A. C. & Moras, D. Glycyl-tRNA synthetase uses a negatively charged pit for specific recognition and activation of glycine. *J. Mol. Biol.* **286**, 1449-1459 (1999).